

ON THE RATE OF CONVERGENCE IN NORMAL APPROXIMATION AND LARGE DEVIATION PROBABILITIES FOR A CLASS OF STATISTICS*

M. L. PURI† AND M. SEOH‡

Summary. A new class of statistics is introduced to include, as special cases, unsigned linear rank statistics, signed linear rank statistics, linear combination of functions of order statistics, linear functions of concomitants of order statistics, and a rank combinatorial statistic. For this class, the rate of convergence to normality and Cramér's type large deviation probabilities are investigated. Under the assumption that underlying observations are only independent, it is shown that this rate is $O(N^{-\delta/2} \log N)$ if the first derivative of the score generating function φ satisfies Lipschitz's condition of order δ , $0 < \delta \leq 1$, and it is $O(N^{-1/2})$ if φ'' satisfies Lipschitz's condition of order $\delta \geq \frac{1}{2}$; and that Cramér's large deviation theorem holds in the optimal range $0 < x \leq \rho_N N^{1/6}$, $\rho_N = o(1)$. The results obtained provide new results and extend as well as generalize a number of known results obtained in this direction.

1. Introduction. Let $Z_{Nj} = (Y_{Nj}^{(1)}, Y_{Nj}^{(2)}, \dots, Y_{Nj}^{(p)})$, $1 \leq j \leq N$, be independent p -variate r.v.'s (random vectors) with joint c.d.f. (cumulative distribution function) $F_{Nj}(t)$, $t \in \mathbf{R}^p$ (the Euclidean p -dimensional space). Let g be a function on \mathbf{R}^p such that $X_{Nj}^g = g(Z_{Nj})$, $1 \leq j \leq N$, has a c.d.f. G_{Nj} which may not be continuous. Define the rank R_{Nj}^g of X_{Nj}^g among $\{X_{Nk}^g; 1 \leq k \leq N\}$ by $R_{Nj}^g = \sum_{k=1}^N u(X_{Nj}^g - X_{Nk}^g)$, where $u(x) = 1$ or 0 according to whether $x \geq 0$ or $x < 0$. Consider a class of statistics

$$(1.1) \quad T_N = \sum_{j=1}^N c_{Nj} a_{NR_{Nj}^g}(Z_{Nj}),$$

where c_{Nj} , $1 \leq j \leq N$, are regression constants and $a_{Nj}(t)$, $1 \leq j \leq N$, are functions (called scores) generated by a known nonconstant function (called a score generating function) $\varphi(s, t)$, $0 < s < 1$, $t \in \mathbf{R}^p$, in either of the following two ways:

$$(1.2) \quad a_{Nj}(t) = \varphi(\mathbf{E} U_{N:j}, t), \quad 1 \leq j \leq N, \quad (\text{approximate scores}),$$

$$(1.3) \quad a_{Nj}(t) = \mathbf{E} \varphi(U_{N:j}, t), \quad 1 \leq j \leq N, \quad (\text{exact scores}).$$

Here $U_{N:j}$ denotes the j th order statistic in a random sample of size N from the uniform distribution on $(0, 1)$.

The statistic (1.1) includes as special cases the following statistics.

(a) *Unsigned linear rank statistic* $\mathcal{S}_1 = \sum_{j=1}^N c_{Nj} a_{NR_{Nj}^*}^*$ by taking $p = 1$, $g(t) = t$ and $\varphi(s, t) = \tilde{\varphi}(s)$.

(b) *Signed linear rank statistic* $\mathcal{S}_2 = \sum_{j=1}^N c_{Nj} a_{NR_{Nj}^+}^* \operatorname{sgn} Y_{Nj}^{(1)}$ by taking $p = 1$, $g(t) = |t|$ and $\varphi(s, t) = \tilde{\varphi}(s) \operatorname{sgn} t$ (where $\operatorname{sgn} t = 2u(t) - 1$).

(c) *Linear combination of functions of order statistics* $\mathcal{S}_3 = (1/N) \sum_{j=1}^N a_{NR_{Nj}}^* \psi(Y_{Nj}^{(1)}) = (1/N) \sum_{j=1}^N a_{Nj}^* \psi(Y_{N:j}^{(1)})$ by taking $p = 1$, $g(t) = t$, $\varphi(s, t) = \tilde{\varphi}(s) \psi(t)$ and $c_{Nj} = N^{-1}$, $1 \leq j \leq N$.

* Received by the editors July 2, 1984

† Department of Mathematics, Indiana University, Bloomington, Indiana 47405. The U.S.A. research was supported by Office of Naval Research contract N14-85-K-0648.

‡ Department of Mathematics and Statistics, Wright State University, Dayton, Ohio 45435, U.S.A.

(d) *Linear functions of concomitants of order statistics* $\mathcal{S}_4 = (1/N) \sum_{j=1}^N a_{Nj}^* \psi(Y_{Nj}^{(2)}) = (1/N) \sum_{j=1}^N a_{Nj}^* \psi(Y_{N:j}^{(2)})$ by taking $p=2$, $g(t_1, t_2) = t_1$, $\varphi(s, t_1, t_2) = \tilde{\varphi}(s)\psi(t_2)$ and $c_{Nj} = N^{-1}$, $1 \leq j \leq N$.

Here the a_{Nj}^* are the usual scores generated by $\tilde{\varphi}(s)$; R_{Nj} , $Y_{N:j}^{(1)}$, $1 \leq j \leq N$, are the rank of $Y_{Nj}^{(1)}$ and the j th order statistic, respectively, among $\{Y_{Nk}^{(1)}, 1 \leq k \leq N\}$; R_{Nj}^+ , $1 \leq j \leq N$, is the rank of $|Y_{Nj}^{(1)}|$ among $\{|Y_{Nk}^{(1)}|, 1 \leq k \leq N\}$; $Y_{N:j}^{(2)}$ is the concomitant of the j th order statistic $Y_{N:j}^{(1)}$ (j th induced order statistic, see [2]), i.e., $Y_{N:j}^{(2)} = Y_{N:R_{Nj}}^{(2)}$, if the marginal c.d.f.'s of $Y_{Nj}^{(1)}$, $1 \leq j \leq N$, are continuous. (Note that the equalities in \mathcal{S}_3 and \mathcal{S}_4 are justified under this condition.) Furthermore, denoting $W_{jk} = c_{Nj}a_{Nk}(Z_{Nj})$, $1 \leq j, k \leq N$, we can rewrite (1.1) as

(e) *Rank combinatorial statistic* $\mathcal{S}_5 = \sum_{j=1}^N W_{jR_{Nj}^+}$ (cf. [25]).

In recent years, under different sets of assumptions, a number of authors have investigated these statistics separately, to derive asymptotic normality, rates of convergence to normality and the Cramér-type large deviation results. See, for example, Hájek [8], [9], Does [4], and Kallenberg [17] for results concerning the statistic \mathcal{S}_1 ; Hušková [13] and Müller-Funk and Witting [19] for results concerning \mathcal{S}_2 ; Stigler [22], Helmers [10], [11], and Vandermale and Veraverbeke [24] for results concerning \mathcal{S}_3 ; Yang [26] for \mathcal{S}_4 ; and Motoo [18] and von Bahr [25] for results concerning \mathcal{S}_5 , among others.

In this paper, we derive suitable order bounds for the remainder term in the normal approximation, and the Cramér-type large deviation probabilities for the class of statistics T_N defined in (1.1). The results obtained include, as special cases, Bergström and Puri [1], Kallenberg [17], Vandermale and Veraverbeke [24], and [27], [29]. They also extend and generalize Hušková [14], [15], Does [4], Müller-Funk and Witting [19], Bjerre [3] and Helmers [10], [11], among others (see Remarks 3.1–3.5 and 4.1–4.4).

2. Assumption and notation. We assume the following.

ASSUMPTION $A_l^{(\delta)}$. The score generating function $\varphi(s, t)$, $0 < s < 1$, $t \in \mathbb{R}^p$, has the l th partial derivatives $\varphi_1^{(l)}(s, t) = \theta^l(s, t)/\partial s^l$ satisfying Lipschitz's condition of order δ , $0 < \delta \leq 1$, with respect to its first argument s , i.e.,

$$(2.1) \quad |\varphi_1^{(l)}(x, t) - \varphi_1^{(l)}(y, t)| \leq |x - y|^\delta \Delta_l(t), \quad 0 < x, y < 1,$$

where Δ_l is some non-negative function. There are non-negative absolute constants K and η such that

$$(2.2) \quad \max_{1 \leq j \leq N} \mathbb{E}|\Delta_l(Z_{Nj})|^r \leq K^r r^{\eta r} \quad \text{for any real } r > 0.$$

Furthermore, $\varphi(s, t)$, $\varphi_1^{(1)}(s, t)$, \dots , $\varphi_1^{(l)}(s, t)$ are jointly measurable in s and t such that

$$(2.3) \quad |\varphi_1^{(k)}(s, t)| \leq \Delta_k(t), \quad k = 0, 1, \dots, l \quad (\varphi_1^{(0)} \equiv \varphi).$$

Remark 2.1. The condition (2.1) ensures that we can choose a suitable function Δ_l satisfying (2.3) without loss of generality.

Remark 2.2. With the special form $\varphi(s, t) = \tilde{\varphi}(s)\psi(t)$ of the score generating function, we note that the joint measurability is out of question and the condition (2.1) restricts the score generating function $\tilde{\varphi}$ to bounded ones.

The main idea (utilized essentially by everyone) is as follows: assuming $\varphi_1(\equiv \varphi_1^{(1)})$, we first approximate the statistic T_N in (1.1) with approximate scores (1.2) by the two-term Taylor expansion, i.e.,

$$(2.4) \quad S_N = \sum_{j=1}^N c_{Nj} \{ \varphi(\rho_{Njj}, Z_{Nj}) + (\rho_{Nj} - \rho_{Njj}) \varphi_1(\rho_{Njj}, Z_{Nj}) \},$$

where $\rho_{Nj} = R_{Nj}^g / (N+1)$ and $\rho_{Njj} = E(\rho_{Nj} | Z_{Nj})$, $1 \leq j \leq N$, and then treat the remainder term under suitable assumptions. Next, we use Hájek's [9] projection method to approximate S_N by

$$(2.5) \quad \hat{S}_N = \sum_{i=1}^N E(S_N | Z_{Ni}) - (N-1)ES_N.$$

Then, the desired properties of T_N can be established by showing those of \hat{S}_N as applications of the well-known classical theory for sums of independent r.v.'s.

Let $\varphi(x)$ denote the standard normal c.d.f. and put

$$(2.6) \quad \tau_N^2 = \mathbf{D}T_N, \quad \sigma_N^2 = \mathbf{D}S_N, \quad \hat{\sigma}_N^2 = \mathbf{D}\hat{S}_N.$$

Furthermore, let $\|\cdot\|$ denote the usual supremum norm and, for any r.v. S , let S^0 denote the r.v. S centered at its expectation, i.e., $S^0 = S - ES$.

From now on, we omit the first subscript N in Z_{Nj} , X_{Nj}^g , c_{Nj} , ρ_{Njj} , etc., for the ease of notation whenever it causes no confusion.

3. The rate of convergence in the normal approximation. The asymptotic normality of the statistics \mathcal{S}_1 and \mathcal{S}_5 (special cases of the statistics T_N in (1.1)) is well known. However, often one needs information more precise than the asymptotic normality can provide, and one may try to find suitable order bounds for the remainder term $\|\mathbf{P}(T_N^0 \leq \tau_N \cdot) - \varphi(\cdot)\|$ when the distribution of (1.1) is approximated by the normal one.

THEOREM 3.1. *If assumption $A_1^{(\delta)}$ is satisfied, then*

$$(3.1) \quad \|\mathbf{P}(T_N^0 \leq \hat{\sigma}_N \cdot) - \Phi(\cdot)\| \leq 36CK^3 3^{3\eta} \omega_3 + 2eM_N (\log N)^{1+\eta_N-\delta/2},$$

$$(3.2) \quad |\tau_N - \hat{\sigma}_N| \leq M_N \hat{\sigma}_N N^{-\delta/2},$$

where C is a constant (0.7975 is van Beek's [23] estimation) and

$$(3.3) \quad M_N = 64eK 2^\eta \hat{\sigma}_N^{-1} \left(\sum_{j=1}^N c_j^2 \right)^{1/2}.$$

THEOREM 3.2. *If assumption $A_2^{(\delta)}$ is satisfied with a bounded Δ_2 , then*

$$(3.4) \quad \|\mathbf{P}(T_N^0 \leq \hat{\sigma}_N \cdot) - \varphi(\cdot)\| = O(\max\{\omega_3, \omega_3^{2\delta}\}) \quad \text{and} \quad |\tau_N - \hat{\sigma}_N| = O(\hat{\sigma}_N \omega_3),$$

where $\omega_3 = \hat{\sigma}_N^{-3} \sum_{j=1}^N |c_j|^3$.

Theorems 3.1 and 3.2 not only extend or generalize a number of existing theorems, they also provide new results as explained in the following remarks. (In what follows the general alternatives mean that the underlying distributions are only independent.)

Remark 3.1 (Unsigned linear rank statistic \mathcal{S}_1). Theorem 3.1 includes Theorem 1.2 in Bergström and Puri [1] that obtained the order bound $N^{-1/2} \log N$ only for approximate scores. Our assumptions are slightly weaker, and our bound is more explicit than the Bergström–Puri bound. As in our Theorem 3.2, Hušková [14], [15] and Does [4] obtained the Berry–Esséen bound of order $N^{-1/2}$ (the former dealt with bounded φ and the latter included unbounded φ) under assumptions on φ weaker than ours. However, their results are derived under the restrictions that underlying distributions are continuous and identical (Hušková, Does) or contiguous (Hušková).

Remark 3.2 (Signed linear rank statistic \mathcal{S}_2). Müller-Funk and Witting [19] derived a bound of order $O(N^{-1/2}(\log N)^2)$ for a restrictive case of statistic \mathcal{S}_2 (assuming that the underlying distributions are i.i.d. and $c_{N1} = \dots = c_{NN}$). For the case of general alternatives, Ralescu and Puri [21] derived a bound of order $O(N^{-\alpha})$, $\alpha > \frac{1}{2}$.

Remark 3.3 (Linear combination of functions of order statistic \mathcal{S}_3). Bjerve [3], dealing with the trimmed linear combination of order statistics, obtained the rate $O(N^{-1/2})$ for the case of i.i.d. r.v.'s imposing rather strong smoothness conditions on the underlying distribution function. Helmers [10], [11], who improved Bjerve's results and considered the linear combination of order statistic only for the approximate scores, obtained the same rate as Bjerve under assumption on φ weaker than ours. However, both Bjerve and Helmers deal with linear combination of order statistics for the i.i.d. r.v.'s while we consider the problem under the general alternative both for approximate as well as exact scores.

Remark 3.4 (Linear functions of concomitants of order statistics \mathcal{S}_4). Our Theorems 3.1 and 3.2 provide rates of convergence to normality for the statistic \mathcal{S}_4 that to the best of our knowledge, has not been investigated so far. This would justify the application in Yang [26].

Remark 3.5 (Rank combinatorial statistic \mathcal{S}_5). Von Bahr [25] considered a class of statistics that includes \mathcal{S}_5 as a special case, and derived the order $N^{-1/2}$ under rather restrictive assumption that the vector of ranks is from the i.i.d. r.v.'s.

4. Cramér's type large deviation probabilities. In this section, we are concerned with relative error in the normal approximation of the distribution of the statistic (1.1), i.e., $\mathbf{P}(T_N^0) > \tau_N x [1 - \varphi(x)]^{-1}$.

We assume that the variance of T_N satisfies

$$(4.1) \quad \underline{\lim} \tau_N^2 = \underline{\lim} DT_N > 0,$$

and that the regression constants satisfy

$$(4.2) \quad \max_{1 \leq j \leq N} |c_j| = O(N^{-1/3}), \quad \sum_{j=1}^N c_j^2 = 1, \quad \sum_{j=1}^N |c_j|^3 = O(N^{-1/2}).$$

Let $\beta = 1/(2(3+2\eta))$. Then, we have

THEOREM 4.1. *If Z_1, Z_2, \dots, Z_N are i.i.d. r.v.'s, $c_1 = c_2 = \dots = c_N$, and assumption $A_1^{(1)}$ is satisfied, then uniformly in the region $0 < x \leq \rho_n N^\beta$, $\rho_n = o(1)$,*

$$(4.3) \quad \mathbf{P}(T_N^0 \geq \tau_N x) [1 - \Phi(x)]^{-1} = 1 + o(1) \quad \text{as } N \rightarrow \infty$$

which remains true if τ_N is replaced by $\hat{\sigma}_N$.

THEOREM 4.2. *If Z_1, Z_2, \dots, Z_N are only independent and assumption $A_1^{(1)}$ is satisfied with a bounded Δ_1 , then the conclusion of Theorem 4.1 holds.*

Theorems 4.1 and 4.2 provide a number of new results, as well as include known results as special cases. For reasons of clarity, we make the following remarks.

Remark 4.1 (Unsigned linear rank statistic \mathcal{S}_1). Kallenberg's [17] result is a special case of our Theorem 4.2. Kallenberg obtained the result under the assumption that underlying distributions are i.i.d. and continuous while ours is obtained under the general alternatives and without assuming the continuity of the underlying distribution functions.

Remark 4.2 (Signed linear rank statistic \mathcal{S}_2). The same result as Kallenberg's is obtained under the general alternative that to the best of our knowledge is not known.

Remark 4.3 (Linear combinations of functions of order statistics \mathcal{S}_3). Our Theorem 4.1 includes, as a special case, Theorem 2 in Vandemaële and Veraverbeke [24]. However, the latter applies only to the linear combinations of order statistics without continuity of underlying distributions while the former applies to the linear combinations of functions of order statistics with the continuity assumption. Our Theorem 4.2 extends these results under the general alternatives; however, our score generating

function is uniformly bounded to cover only the linear combination of bounded functions of order statistics.

Remark 4.4 (Linear functions of concomitants of order statistics \mathcal{S}_4 and rank combinatorial statistic \mathcal{S}_5). Theorems 4.1 and 4.2 provide, to the best of our knowledge, the only known results concerning large deviation probabilities both for the statistics \mathcal{S}_4 and \mathcal{S}_5 .

5. Preliminaries. In this section, we derive bounds on the 2rth moments for the r.v.'s defined in the following lemmas. These bounds play important roles in the proofs of our theorems.

Let $\{Y_j\}_{j=1}^\infty$ be a sequence of r.v.'s and $\{d_j\}$ a sequence of real numbers.

LEMMA 5.1. Let $W_j, j \geq 1$, be r.v.'s of the form $W_j = g_j(Y_1, Y_2, \dots, Y_j)$ such that $E(W_j | Y_1, Y_2, \dots, Y_{j-1}) = 0$ for $j \geq 2$. If the sequence $\{d_j\}$ is nonincreasing in absolute values, then for any positive integers r and l ,

$$(5.1) \quad E\left(\sum_{j=1}^l d_j W_j\right)^{2r} \leq (4e)^r \left(\sum_{j=1}^l d_j^2\right)^r r^r \max_{1 \leq j \leq l} E W_j^{2r}.$$

LEMMA 5.2. Let $Y_j, j \geq 1$, be independent and let \tilde{V}_{jk} be r.v.'s of the form $\tilde{V}_{jk} = g_{jk}(Y_j, Y_k)$, $1 \leq j, k < \infty$, such that for any j and $k, j \neq k$, $E(\tilde{V}_{jk} | Y_j) = E(\tilde{V}_{jk} | Y_k) = 0$. Then, for any positive integers l and r

$$(5.2) \quad E\left(\frac{1}{l} \sum_{j=1}^l \sum_{k \neq j}^l d_j \tilde{V}_{jk}\right)^{2r} \leq (4e)^{2r} \left(\sum_{j=1}^l d_j^2\right)^r (2r)^{2r} l^{-r} \max_{\substack{1 \leq j, k \leq l \\ j \neq k}} E \tilde{V}_{jk}^{2r}.$$

Proof of Lemma 5.1. For $r \geq l$, the proof follows by applying Hölder's inequality to $(\sum_{j=1}^l d_j W_j)^{2r}$ and some routine computations. For $r < l$, the proof follows by induction on l with r fixed.

Proof of Lemma 5.2. Since the assumptions and the conclusion of this lemma are invariant under simultaneous permutation of the d_j and Y_j , we may assume that $|d_1| \geq |d_2| \geq \dots \geq |d_l|$, without loss of generality.

Define $W_1 = \tilde{W}_1 = 0$, $W_j = \sum_{k=1}^{j-1} \tilde{V}_{jk}$ and $\tilde{W}_j = \sum_{k=1}^{j-1} d_k \tilde{V}_{kj}$, $2 \leq j \leq l$, so that

$$\hat{V}_l = \frac{1}{l} \sum_{j=1}^l \sum_{k \neq j}^l d_j \tilde{V}_{jk} = \frac{1}{l} \left\{ \sum_{j=1}^l d_j W_j + \sum_{j=1}^l \tilde{W}_j \right\}.$$

The proof follows by noting that $\{W_j\}_{j=1}^l$ and $\{\tilde{W}_j\}_{j=1}^l$ satisfy the assumptions of Lemma 5.1, and using the following facts:

$$\begin{aligned} (a) \quad E \hat{V}_l^{2r} &\leq l^{-2r} 2^{2r-1} \left\{ E \left(\sum_{j=1}^l d_j W_j \right)^{2r} + E \left(\sum_{j=1}^l \tilde{W}_j \right)^{2r} \right\} \\ &\leq l^{-2r} 2^{2r-1} (4e)^r r^r \left\{ \left(\sum_{j=1}^l d_j^2 \right)^r \max_{1 \leq j \leq l} E W_j^{2r} + l^r \max_{1 \leq j \leq l} E \tilde{W}_j^{2r} \right\}, \\ (b) \quad E W_j^{2r} &\leq (4e)^r (j-1)^r r^r \max_{\substack{1 \leq k \leq l \\ k \neq j}} E \tilde{V}_{jk}^{2r}, \\ (c) \quad E \tilde{W}_j^{2r} &\leq (4e)^r \left(\sum_{k=1}^{j-1} d_k^2 \right)^r r^r \max_{\substack{1 \leq k \leq l \\ k \neq j}} E \tilde{V}_{kj}^{2r}, \end{aligned}$$

which can easily be obtained by repeated use of Lemma 5.1.

We now derive suitable bounds for remainders in approximations of S_N by \hat{S}_N and of T_N by S_N . Note that

$$(5.3) \quad \rho_j - \rho_{jj} = \frac{1}{N+1} \sum_{k \neq j}^N \{u(X_j^g - X_k^g) - G_k(X_j^g)\} \equiv \frac{1}{N+1} \sum_{k \neq j}^N H_{jk}.$$

For each j , conditionally given Z_j , $\rho_j - \rho_{jj}$ is the sum of independent r.v.'s with zero means. Thus an application of Lemma 5.1 and Hölder's inequality yield that, for any real r ,

$$(5.4) \quad \mathbf{E}\{(\rho_j - \rho_{jj})^{2r} | Z_j\} \leq (4e)^r [r]^r N^{-r}, \quad j = 1, 2, \dots, N,$$

where $[r]$ is the smallest integer $\geq r$. It follows by elementary computations that

$$\begin{aligned} \mathbf{E}(S_N | Z_l) &= c_l \varphi(\rho_{ll}, Z_l) + \mathbf{E}S_N - c_l \mathbf{E}\varphi(\rho_{ll}, Z_l) \\ &\quad + (N+1)^{-1} \sum_{j \neq l}^N c_j \mathbf{E}\{H_{jl} \varphi_1(\rho_{jj}, Z_j) | Z_l\}, \quad 1 \leq l \leq N. \end{aligned}$$

Thus we can write

$$(5.5) \quad \hat{S}_N^0 = \sum_{j=1}^N \hat{S}_N^{(j)}, \quad S_N^0 - \hat{S}_N^0 = \frac{1}{N+1} \sum_{j=1}^N \sum_{k \neq j}^N c_j V_{jk},$$

where

$$(5.6) \quad \begin{aligned} \hat{S}_N^{(j)} &= c_j \{\varphi(\rho_{jj}, Z_j) - \mathbf{E}\varphi(\rho_{jj}, Z_j)\} + \frac{1}{N+1} \sum_{k \neq j}^N c_k \mathbf{E}\{H_{kj} \varphi_1(\rho_{kk}, Z_k) | Z_j\}, \\ V_{jk} &= H_{jk} \varphi_1(\rho_{jj}, Z_j) - \mathbf{E}\{H_{jk} \varphi_1(\rho_{jj}, Z_j) | Z_k\}, \end{aligned}$$

and H_{jk} is defined in (5.3). Since $\mathbf{E}\{V_{jk} | Z_j\} = \mathbf{E}\{V_{jk} | Z_k\} = 0$ for $j \neq k$, we may apply Lemma 5.2 to obtain that, for any integer $r \geq 1$,

$$(5.7) \quad \mathbf{E}(S_N^0 - \hat{S}_N^0)^{2r} \leq (8e)^{2r} \left(\sum_{j=1}^N c_j^2 \right)^r (2r)^{2r} N^{-r} \max_{1 \leq j \leq N} \mathbf{E}\Delta_1^{2r}(Z_j).$$

We finally prove

LEMMA 5.3. For any integer $r > 0$, we have, under assumption $A_1^{(\delta)}$,

$$(5.8) \quad \mathbf{E}(T_N^0 - S_N^0)^{2r} \leq (16eK)^{2r} \left(\sum_{j=1}^N c_j^2 \right)^r (2r)^{(1+\eta)2r} N^{-\delta r},$$

under assumption $A_l^{(\delta)}$, $l \geq 1$ (for approximate scores),

$$(5.9) \quad \mathbf{E} \left(T_N^0 - \sum_{k=0}^l I_k^0 \right)^{2r} \leq \left(\frac{2K}{l!} \right)^{2r} \left(\sum_{j=1}^N c_j^2 \right)^r (4er(l+1))^{r(l+\delta)} (2r)^{2r\eta_N - r(l+\delta-1)},$$

where the r.v.'s I_k , $0 \leq k \leq l$, are given by

$$(5.10) \quad I_k = \sum_{j=1}^N c_j \frac{(\rho_j - \rho_{jj})^k}{k!} \varphi_1^{(k)}(\rho_{jj}, Z_j);$$

and, under assumption $A_l^{(\delta)}$, $l \geq 2$,

$$(5.11) \quad \mathbf{E} \left(\sum_{j=1}^N c_j \left\{ \tilde{a}_{R_j}(Z_j) - \varphi \left(\frac{R_j}{N+1}, Z_j \right) \right\}^0 \right)^{2r} \leq \left(\sum_{j=1}^N c_j^2 \right)^r N^{-r} \max_{1 \leq j \leq N} \mathbf{E}\Delta_2^{2r}(Z_j),$$

where $\tilde{a}_j(t) = \mathbf{E}\varphi(U_{N:j}, t)$, $1 \leq j \leq N$, $t \in \mathbf{R}^p$.

Proof. Under assumption $A_l^{(\delta)}$, Taylor's expansion yields that

$$(5.12) \quad \left| \varphi(x, t) - \sum_{k=0}^l \frac{(x-y)^k}{k!} \varphi_1^{(k)}(y, t) \right| \leq \frac{1}{l!} |x-y|^{l+\delta} \Delta_l(t)$$

for any $0 < x, y < 1$ and $t \in \mathbf{R}^p$.

Let \tilde{T}_N denote the statistic (1.1) with exact scores (given by (1.3)) to distinguish it from the statistic T_N with approximate scores. Then it follows from (5.4), (5.12), and Hölder's inequality that

$$(5.13) \quad \begin{aligned} \mathbf{E} \left(T_N^0 - \sum_{k=0}^l I_k^0 \right)^{2r} &\leq 2^{2r} \mathbf{E} \left(\sum_{j=1}^N c_j \left\{ \varphi(\rho_j, Z_j) - \sum_{k=0}^l \frac{(\rho_j - \rho_{jj})^k}{k!} \varphi_1^{(k)}(\rho_{jj}, Z_j) \right\} \right)^{2r} \\ &\leq 2^{2r} \left(\sum_{j=1}^N c_j^2 \right)^r (l!)^{-2r} N^{r-1} \sum_{j=1}^N \mathbf{E} |\rho_j - \rho_{jj}|^{2r(l+\delta)} \Delta_l^{2r}(Z_j) \\ &\leq \left(\frac{2}{l!} \right)^{2r} \left(\sum_{j=1}^N c_j^2 \right)^r (4e)^{r(l+\delta)} \\ &\quad \cdot (r(l+1))^{r(l+\delta)} N^{-r(l+\delta-1)} \max_{1 \leq j \leq N} \mathbf{E} \Delta_l^{2r}(Z_j) \end{aligned}$$

which, together with (2.2), yields (5.9). Moreover,

$$(5.14) \quad \mathbf{E}(\tilde{T}_N^0 - T_N^0)^{2r} \leq 2^{2r} \left(\sum_{i=1}^N c_i^2 \right)^r \mathbf{E} \left[\sum_{j=1}^N \left\{ \bar{a}_{R_j}(Z_j) - \varphi \left(\frac{R_j}{N+1}, Z_j \right) \right\}^2 \right]^r.$$

Under assumption $A_1^{(\delta)}$, (5.12) ensures that $|\tilde{a}_j(t) - \varphi(i/(N+1), t)| \leq N^{-(1+\delta)/2} \Delta_1(t)$ for any $1 \leq j \leq N$ and t . Since $S_N = I_0 + I_1$, (5.8) follows from (5.13) and (5.14). On the other hand, under assumption $A_l^{(\delta)}$, $l \geq 2$, there are second partial derivatives such that $|\tilde{a}_j(t) - \varphi(i/(N+1), t)| \leq (2N)^{-1} \Delta_2(t)$ for any $1 \leq j \leq N$ and t . Hence (5.11) follows from (5.14). The proof follows.

6. Proof of theorems in § 3. We now prove Theorems 3.1 and 3.2.

For convenience, we denote, for any integer r , $\omega_r = \sum_{j=1}^N |c_j^*|^r$, where $c_j^* = c_j \hat{\sigma}_N^{-1}$, $1 \leq j \leq N$.

Proof of Theorem 3.1. By standard arguments, we have, for any $\varepsilon > 0$,

$$(6.1) \quad \begin{aligned} \mathbf{P}(T_N^0 \leq \hat{\sigma}_N x) &\leq \mathbf{P}(\hat{S}_N^0 \leq \hat{\sigma}_N(x + \varepsilon)) + \mathbf{P}(|T_N^0 - \hat{S}_N^0| \geq \hat{\sigma}_N \varepsilon), \\ \mathbf{P}(T_N^0 \leq \hat{\sigma}_N x) &\geq \mathbf{P}(\hat{S}_N^0 \leq \hat{\sigma}_N(x - \varepsilon)) - \mathbf{P}(|T_N^0 - \hat{S}_N^0| \geq \hat{\sigma}_N \varepsilon). \end{aligned}$$

Since $\hat{\sigma}_N^{-1} \hat{S}_N^0 = \sum_{j=1}^N \hat{\sigma}_N^{-1} \hat{S}_N^{(j)}$ (cf. (5.5)) is a normalized sum of independent r.v.'s and since, because of (2.2),

$$(6.2) \quad \sum_{j=1}^N \mathbf{E} |\hat{\sigma}_N^{-1} \hat{S}_N^{(j)}|^3 \leq 4 \sum_{j=1}^N \{8|c_j^*|^3 + N^{-1} \omega_3\} \max_{1 \leq j \leq N} \mathbf{E} \Delta_1^3(Z_j) \leq L_N,$$

where $L_N = 36K^3 3^{3\eta} \omega_3$, we may apply Lemma V.2.1 and Theorem V.2.3 of Petrov [20] to obtain that

$$(6.3) \quad |\mathbf{E} \exp(it \hat{\sigma}_N^{-1} \hat{S}_N^0) - e^{-t^2/2}| \leq 16L_N |t|^3 e^{-t^2/3}, \quad |t| \leq (4L_N)^{-1}$$

and

$$(6.4) \quad \|\mathbf{P}(\hat{S}_N^0 \leq \hat{\sigma}_N \cdot) - \Phi(\cdot)\| \leq CL_N$$

for some absolute constant C (van Beek's [23] estimation is 0.7975).

It follows from (2.2), (5.7), and (5.8) that

$$(6.5) \quad \mathbf{E}(T_N^0 - \hat{S}_N^0)^{2r} \leq (64eK2^\eta)^{2r} \left(\sum_{j=1}^N c_j^2 \right)^r r^{2r(1+\eta)} N^{-\delta r}.$$

Using the well-known inequality $|\Phi(x+y) - \Phi(x)| \leq |y|$, we derive from (6.1), (6.4), and (6.5),

$$(6.6) \quad \|\mathbf{P}(T_N^0 \leq \hat{\sigma}_N \cdot) - \Phi(\cdot)\| \leq CL_N + \varepsilon + (\hat{\sigma}_N \varepsilon)^{-2r} B_N^{(r)},$$

where

$$B_N^{(r)} = \left[64eK2^\eta \left(\sum_{j=1}^N c_j^2 \right)^{1/2} r^{1+n} N^{-\delta/2} \right]^{2r}.$$

Now choose $\varepsilon = (\hat{\alpha}_N \varepsilon)^{-2r} B_N^{(r)}$, i.e., $\varepsilon \leq M_N (r^{1+\eta} N^{-\delta/2})^{2r/(2r+1)}$, where M_N is given by (3.3) and take $r = \lceil -1 + \log N \rceil$ so that $\varepsilon \leq M_N e (\log N)^{1+\eta} N^{-\delta/2}$. This inequality and (6.6) ensure (3.1), and (6.5) yields (3.2). The proof follows.

Proof of Theorem 3.2. The theorem is trivial if $\omega_3 \geq e^{-2}$. Thus we may, without loss of generality, assume that

$$(6.7) \quad \omega_3 < e^{-2}, \quad \|\Delta_2\| \leq \frac{1}{6},$$

which ensures

$$(6.8) \quad \mathbf{E}(\hat{\sigma}_N^{-1} \hat{S}_N^{(j)})^2 \leq ((c_i^*)^2 + N^{-1} \omega_2)/18, \quad \mathbf{E}|\hat{\sigma}_N^{-1} S_N^{(j)}|^3 \leq (8|c_j^*|^3 + N^{-1} \omega_3)/54.$$

Moreover, because $1 = \sum_{j=1}^N \mathbf{E}(\hat{\sigma}_N^{-1} \hat{S}_N^{(j)})^2 \leq 9^{-1} \omega_2$, Hölder's inequality yields

$$(6.9) \quad N^{-1/2} \leq 3^{-3} \omega_3, \quad \omega_1 \leq N^{2/3} \omega_3^{1/3}, \quad \omega_2 \leq N^{1/3} \omega_3^{2/3}.$$

We first prove the theorem for approximate scores.

By standard arguments similar to (6.1), to prove the theorem, it suffices to show that

$$(6.10) \quad \mathbf{P}(|T_N^0 - S_N^0 - I_2^0 - I_{21}^0| \geq \hat{\sigma}_N \omega_3) = O(\max\{\omega_3, \omega_3^{2\delta}\}),$$

$$(6.11) \quad \|\mathbf{P}(S_N^0 + I_{22}^0 \leq \hat{\sigma}_N \cdot) - \Phi(\cdot)\| = O(\omega_3),$$

where I_2 is defined by (5.10) and I_{21} and I_{22} are given by

$$I_{21} = (2(N+1)^2)^{-1} \sum_{(j,k) \neq} \sum c_j H_{jk}^2 \varphi_1^{(2)}(\rho_{jj}, Z_j),$$

$$I_{22} = (2(N+1)^2)^{-1} \sum_{(j,k,l) \neq} \sum \sum c_j H_{jk} H_{jl} \varphi_1^{(2)}(\rho_{jj}, Z_j),$$

where $(j, k) \neq$ and $(j, k, l) \neq$ mean that summations are taken for different indices j , k , and l over the set $\{1, 2, \dots, N\}$. It follows that $I_2 = I_{21} + I_{22}$.

Since $S_N = I_0 + I_1$, applying (5.9) with $r = 1$, $l = 2$, and then using (6.9), we obtain $\mathbf{E}(\hat{\sigma}_N^{-1}(T_N^0 - S_N^0 - I_2^0))^2 = O(\omega_3^{2+2\delta})$. Also direct computations yield $\mathbf{E}(\hat{\sigma}_N^{-1} I_{21}^0)^2 = O(\omega_3^4)$. Equation (6.10) now follows easily by Chebyshev's inequality.

It follows by Esseen's smoothing lemma (see, e.g., Feller [7, p. 603] and (6.3) that, to prove (6.11), it suffices to show that, for some positive $\varepsilon \leq (144K^3 3^{3\eta})^{-1}$,

$$(6.12) \quad \int_{|t| \leq \varepsilon \omega_3^{-1}} |t|^{-1} |f(t)| dt = O(\omega_3),$$

where $f(t) = \mathbf{E} \exp \{it \hat{\sigma}_N^{-1} (S_N^0 + I_{22}^0)\} - \mathbf{E} \exp \{it \hat{\sigma}_N^{-1} \hat{S}_N^0\}$.

We now need two lemmas that are extremely crucial and the proofs of which are given in the Appendix.

LEMMA 6.1. *Under (6.7), we have*

$$\mathbf{E}(\hat{\sigma}_N^{-1} I_{22}^0)^{2r} \leq (2e \|\Delta_2\|/9)^{2r} (2r)^{2r} \omega_3^{2r}.$$

LEMMA 6.2. Under (6.7), there is an absolute constant K such that

$$|\mathbf{E} \exp(it\hat{\sigma}_N^{-1}\hat{S}_N^0)\{\hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\}^r| \leq K^r r^{3r} \omega_3(1+|t|+t^2)^r e^{-t^2/5}$$

for $|t| \leq \omega_3^{-1}$ and any positive integer $r \leq 2^{-1} \log \omega_3^{-1}$.

Now for any positive integer k , Taylor's expansion yields

$$(6.13) \quad \left| \frac{f(t)}{t} \right| \leq \sum_{r=1}^{2k-1} \frac{|t|^{r-1}}{r!} |\mathbf{E}[\{\hat{\alpha}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\}^r \exp(it\hat{\sigma}_N^{-1}\hat{S}_N^0)]| \\ + O\left(\frac{|t|^{2k-1}}{(2k)!} \mathbf{E}\{\hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\}^{2k}\right).$$

Taking $k=1$, and using (5.7), (6.9), Lemmas 6.1 and 6.2, we see that, for any $\varepsilon_1 > 0$,

$$(6.14) \quad \int_{|t| \leq \varepsilon_1 \omega_3^{-1/2}} \left| \frac{f(t)}{t} \right| dt = O(\omega_3).$$

Next, take $2k = [\log \omega_3^{-1}]$. Lemma 6.2 ensures that for any ε_1 , $0 < \varepsilon_1 \leq 1$,

$$(6.15) \quad \int_{\varepsilon_1 \omega_3^{-1/2} \leq |t| \leq \varepsilon_1 \omega_3^{-1}} \sum_{r=1}^{2k-1} \frac{|t|^{r-1}}{r!} |\mathbf{E}\{\hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\}^r \exp(it\hat{\sigma}_N^{-1}\hat{S}_N^0)| dt \\ \leq \omega_3 \exp\left(-\frac{\varepsilon_1^2 \omega_3^{-1}}{5}\right) (2k)^{6k} \omega_3^{-6k} e^{3K} = O(\omega_3).$$

It also follows from (5.7), (6.9), and Lemma 6.1 that

$$(6.16) \quad \int_{|t| \leq \varepsilon_1 \omega_3^{-1}} \frac{|t|^{2k-1}}{(2k)!} \mathbf{E}\{\hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\}^{2k} dt \leq \frac{(8e\|\Delta_2\|)^{2k} (2k)^{2k} \omega_3^{2k}}{(2k)!} (\varepsilon_1 \omega_3^{-1})^{2k} \\ \leq (2e\varepsilon_1)^{2k} (2k)^{2k} [(2k)!]^{-1} \\ \leq (2e^2 \varepsilon_1)^{2k},$$

where the last equality follows by Stirling's formula (see, e.g., Feller [6]). By taking $\varepsilon_1 = 1/(2e^3)$ and combining (6.13), (6.15), and (6.16), we obtain

$$(6.17) \quad \int_{\varepsilon_1 \omega_3^{-1/2} \leq |t| \leq \varepsilon_1 \omega_3^{-1}} \left| \frac{f(t)}{t} \right| dt = O(\omega_3),$$

which, together with (6.14), ensures (6.12), and consequently (6.11).

Since, because of (5.11),

$$\mathbf{P}\left(\left|\sum_{j=1}^N c_j(\tilde{\alpha}_{R_j}(Z_j) - \varphi(R_j/(N+1)))\right|^0 \geq \hat{\sigma}_N^{-1} \omega_3\right) \leq \omega_3,$$

the proof for exact scores follows analogously. The remaining half of (3.4) follows from (5.7), (5.8), and (6.9). The proof is complete.

7. Proofs of theorems in § 4. Before proving Theorems 4.1 and 4.2, we state some preliminary results. First note that it follows from (2.2), (4.2), (5.7), (5.8), and Hölder's inequality that, for any $r \geq \frac{1}{2}$,

$$(7.1) \quad \mathbf{E}(T_N^0 - \hat{S}_N^0)^{2r} \leq (32eK)^{2r} (2r)^{2r(1+\eta)} N^{-r}.$$

Furthermore, since $|\hat{\sigma}_N^2 - \tau_N^2| \leq 2\tau_N(\mathbf{D}(\hat{S}_N - T_N))^{1/2} + \mathbf{D}(\hat{S}_N - T_N)$, we have

$$(7.2) \quad \hat{\sigma}_N^2 / \tau_N^2 = 1 + O(N^{-1/2}), \quad \hat{\sigma}_N^2 \geq \sigma^2 \quad \text{for some constant } \sigma > 0.$$

Hence, to prove Theorems 4.1 and 4.2, it suffices to show that

$$(7.3) \quad \mathbf{P}(T_N^0 > \hat{\sigma}_N x) [1 - \varphi(x)]^{-1} = 1 + o(1)$$

uniformly in the region $0 < x \leq \rho_N N^\beta$ (recall that $\beta = (2(3 + 2\eta))^{-1}$).

We now consider the standard arguments

$$(7.4) \quad \begin{aligned} \mathbf{P}(T_N^0 > \hat{\sigma}_N x) &\leq \mathbf{P}(\hat{S}_N^0 > \hat{\sigma}_N(x - N^{-\beta})) + \mathbf{P}(|T_N^0 - \hat{S}_N^0| > \hat{\sigma}_N N^{-\beta}), \\ \mathbf{P}(T_N^0 > \hat{\sigma}_N x) &\geq \mathbf{P}(\hat{S}_N^0 > \hat{\sigma}_N(x + N^{-\beta})) - \mathbf{P}(|T_N^0 - \hat{S}_N^0| > \hat{\sigma}_N N^{-\beta}). \end{aligned}$$

Using Chebyshev's inequality and applying (7.1) with $2r = \delta N^{2\beta}$, $\delta = (\sigma/32e^2 K)^{1/(1+\eta)}$, we have

$$\begin{aligned} \mathbf{P}(|T_N^0 - \hat{S}_N^0| > \hat{\sigma}_N N^{-\beta}) &\leq \mathbf{E}(T_N^0 - \hat{S}_N^0)^{2r} N^{2r\beta} \hat{\sigma}_N^{-2r} \\ &\leq \{(32eK/\sigma)(2r)^{1+\eta} N^{-(1+\eta)2\beta}\}^{2r} = \exp(-\delta N^{2\beta}). \end{aligned}$$

Hence, using Lemma VII.1.2 of Feller [6], we derive

$$(7.5) \quad \mathbf{P}(|T_N^0 - \hat{S}_N^0| > \hat{\sigma}_N N^{-\beta}) [1 - \varphi(x)]^{-1} \leq \exp(-\delta N^{2\beta}) [1 - \varphi(\rho_N N^\beta)]^{-1} = o(1)$$

uniformly for $0 < x \leq \rho_N N^\beta$.

Thus, in view of (7.4) and (7.5), to prove (7.3), it suffices to show that

$$(7.6) \quad \mathbf{P}(\hat{S}_N^0 > \hat{\sigma}_N x_N) [1 - \varphi(x)]^{-1} = 1 + o(1), \quad |x_N - x| = N^{-\beta},$$

uniformly in $0 < x \leq \rho_N N^\beta$.

Proof of Theorem 4.1. Define a function

$$h(x) = \int_{\mathbf{R}^p} \{u(g(\underline{t}) - x) - G_j(g(\underline{t}))\} \varphi_1(\Gamma(\underline{t}), \underline{t}) dF(\underline{t}),$$

where $\Gamma(\underline{t}) = \mathbf{E}(\rho_j | Z_j = \underline{t})$ and $F(\underline{t})$ is the common distribution function of the Z_j . Set

$$\tilde{S}_N = \sum_{j=1}^N \hat{S}_N^{(j)} = \sum_{j=1}^N \{\varphi(\rho_{ij}, Z_j) - \mathbf{E}\varphi(\rho_{ij}, Z_j) + h(X_j)\},$$

and denote $\tilde{\sigma}^2 = \text{Var}(\tilde{S}_N^{(j)})$, $1 \leq j \leq N$. Since $\hat{S}_N^0 - N^{-1/2} \tilde{S}_N = -2N^{-1/2}(N+1)^{-1} \sum_{j=1}^N h(X_j)$, it suffices, using the same arguments as above, to show that, for $|x_N - x| = N^{-\beta}$,

$$(7.7) \quad \mathbf{P}(\tilde{S}_N > N^{1/2} \tilde{\sigma} x_N) [1 - \varphi(x)]^{-1} = 1 + o(1)$$

uniformly in the region $0 < x \leq \rho_N N^\beta$. Note that \tilde{S}_N is the sum of i.i.d. r.v.'s with zero means and finite variance $\tilde{\sigma}^2$ such that, with $k_0 = \max\{k: 4\beta k/(2\beta + 1) < 1\}$,

$$\begin{aligned} \mathbf{E} \exp |\tilde{S}_N^{(1)}|^{4\beta/(2\beta+1)} &\leq 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \mathbf{E}(\Delta_1(Z_1) + 2\mathbf{E}\Delta_1(z_1))^{4\beta k/(2\beta+1)} \\ &\leq 1 + \sum_{k=1}^{k_0} \frac{(3K)^{4\beta k/(2\beta+1)}}{k!} + \sum_{k=k_0+1}^{\infty} \frac{16^k}{k!} \left(K \left(\frac{4\beta k}{2\beta+1} \right)^\eta \right)^{4\beta k/(2\beta+1)} \\ &\leq 1 + e^{\tilde{C}} \sum_{k=1}^{\infty} \frac{\tilde{C}^k}{k!} k^{(1-2/(\eta+2))k} < \infty, \end{aligned}$$

where \tilde{C} is an absolute constant; (the last inequality follows by Sterling's formula). Hence (7.7) follows using Petrov [20, p. 251]. The proof follows.

Proof of Theorem 4.2. The method to prove (7.6) is similar to that of Kallenberg [17].

First note that $\hat{S}_N^0 = \sum_{j=1}^N \hat{S}_M^{(j)}$ is the sum of independent r.v.'s with zero means, and

$$|\hat{S}_N^{(j)}| \leq 2|c_j| \|\Delta_1\| + (N+1)^{-1} \sum_{k \neq j}^N |c_k| \|\Delta_1\| \leq 3\|\Delta_1\| \leq 3\|\Delta_1\| A_1 N^{-1/3}, \quad 1 \leq j \leq N,$$

for some constant A_1 . Furthermore there is an integer N_0 such that, for all $N \geq N_0$ and $0 < x < \rho_N N^{1/6}$, $|x_N - x| \leq N^{-1/6}$,

$$(7.8) \quad 0 < 3\|\Delta_1\| A_1 N^{-1/3} \hat{\sigma}_N^{-1} x_N \leq \frac{1}{12}$$

in view of (7.2). Thus we can use Theorem 1 of Feller [5] (see also Petrov [20, p. 253]), to obtain that for all $N \geq N_0$,

$$(7.9) \quad \mathbf{P}(\hat{S}_N^0 > x_N \hat{\sigma}_N) = \exp \left\{ -\frac{1}{2} x_N^2 Q_N(x_N) \right\} [1 - \Phi(x_N) + \theta_N \lambda_N e^{-\lambda_N^2/2}],$$

where

$$(7.10) \quad \lambda_N = 3\|\Delta_1\| A_1 N^{-1/3} \hat{\sigma}_N^{-1}, \quad \theta_N < 7.465; \quad Q_N(x) = \sum_{j=1}^{\infty} q_{Nj} x^j,$$

$$q_{N1} = 3^{-1} \hat{\sigma}_N^{-3} \sum_{j=1}^N \mathbf{E}(\hat{S}_N^{(j)})^3, \quad |q_{Nj}| < 8^{-1} (12\lambda_N)^j, \quad j \geq 2.$$

Note that $|x_N| \leq \rho_N N^{1/6} + N^{-1/6}$. Since $\rho_N = o(1)$, we have $x_N = o(N^{1/6})$ as $N \rightarrow \infty$. Let $K > 0$ be such that

$$(7.11) \quad |x_N| \leq K \hat{\sigma}_N N^{1/6} / (36 A_1 \|\Delta_1\|), \quad N \geq 1.$$

Also it follows by (7.2) and assumption (4.2) that

$$(7.12) \quad |q_{N1} x_N| \leq 4 \hat{\sigma}_N^{-3} \|\Delta_1\| \sum_{j=1}^N |c_j|^3 x_N \leq A_3 N^{-1/3},$$

where A_3 is an absolute constant. Hence, combining (7.10), (7.11), and (7.12), we obtain

$$|Q_N(x_N)| \leq A_3 N^{-1/3} + \sum_{j=2}^{\infty} |q_{Nj} x_N^j| \leq A_3 N^{-1/3} + 8^{-1} \sum_{j=2}^{\infty} (K N^{-1/6})^j,$$

which implies that, as $N \rightarrow \infty$,

$$(7.13) \quad x_N^2 Q_N(x_N) = x_N^2 O(N^{-1/3}) = o(1) \quad \text{uniformly in } 0 < x \leq \rho_N N^{1/6}.$$

Moreover, Lemma VII.1.2 of Feller [6] ensures that

$$(7.14) \quad \theta_N \lambda_N \exp(-\lambda_N^2/2) [1 - \Phi(x_N)]^{-1} = O(1),$$

$$[1 - \Phi(x_N)]^{-1} [1 - \Phi(x)] = 1 + o(1).$$

The proof follows using (7.9), (7.13), and (7.14).

Appendix. We now prove Lemmas 6.1 and 6.2.

Proof of Lemma 6.1. Set $\hat{\sigma}_N^{-1} I_{22}^0 = (2(N+1))^{-1} \sum_{j=1}^N c_j^* \hat{H}_j$, where

$$\hat{H}_j = (N+1)^{-1} \sum_{k \neq j} \sum_{l \neq j, k} H_{jk} H_{jl} \varphi_1^{(2)}(\rho_{jl}, Z_j).$$

Then, an application of Lemma 5.2 yields that, for $1 \leq j \leq N$,

$$\mathbf{E} \hat{H}_j^{2r} \leq (4e \|\Delta_2\|)^{2r} (2r)^{2r}$$

which, together with (6.9) and Hölder's inequality, gives the desired result.

Proof of Lemma 6.2. Since $\log \omega_3^{-1} \leq 2(\max_{1 \leq j \leq N} |c_j^*|)^{-1}$ when $\omega_3 \leq 1$, it suffices to show that

$$(A.1) \quad \begin{aligned} & |\mathbf{E} \exp(it\hat{\sigma}_N^{-1}\hat{S}_N^0)\{\hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\}^r| \\ & \leq K^r r^{3r} \omega_3(1+|t|+t^2)^r \exp\left(\frac{-7t^2}{18} + \frac{t^2 r}{6} \max_{1 \leq j \leq N} |c_j^*|\right). \end{aligned}$$

Let $\mu_j(t)$ be the characteristic function of $\hat{\sigma}_N^{-1}\hat{S}_N^{(j)}$. Then it follows (cf. Petrov [20, p. 110]) that

$$|\mu_j(t)| \leq \exp\left\{-\frac{t^2}{2} \mathbf{E}(\hat{\sigma}_N^{-1}\hat{S}_N^{(j)})^2 + \frac{2}{3}|t|^3 \mathbf{E}|\hat{\sigma}_N^{-1}\hat{S}_N^{(j)}|^3\right\},$$

which, for $|t| \leq \omega_3^{-1}$, yields

$$(A.2) \quad \prod_{j \in J} |\mu_j(t)| \leq \exp\left\{\frac{-7t^2}{18} \frac{t^2}{18} \#J \max_{1 \leq j \leq N} |c_j^*|\right\},$$

where J is a subset of $\{1, 2, \dots, N\}$ and $\#J$ is its cardinality.

Recalling that H_{jk} and V_{jk} , $1 \leq j, k \leq N$, are given by (5.3) and (5.6), respectively, we define r.v.'s for $1 \leq j, k, l \leq N$:

$$(A.3) \quad \tilde{V}_{jkl} = c_j^* \left\{ \frac{N+1}{N-2} V_{jk} + \frac{1}{2} H_{jk} H_{jl} \varphi_1^{(2)}(\rho_{ij}, Z_j) \right\}.$$

Then it follows from (5.5) and the definition of I_{22} that

$$(A.4) \quad \hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0) = (N+1)^{-2} \sum_{(j,k,l) \neq} \sum \sum \tilde{V}_{jkl}.$$

Note that for any distinct indices j, k , and l

$$(A.5) \quad \mathbf{E}(\tilde{V}_{jkl} | Z_j) = \mathbf{E}(\tilde{V}_{jkl} | Z_k) = \mathbf{E}(\tilde{V}_{jkl} | Z_l) = 0.$$

In view of (5.5) and (5.6), $\hat{\sigma}_N^{-1}\hat{S}_N^0$ is a sum of independent r.v.'s. This fact and (A.4) ensure that

$$(A.6) \quad \begin{aligned} & \mathbf{E}\{\exp(it\hat{\sigma}_N^{-1}\hat{S}_N^0)\hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\} \\ & = (N+1)^{-2} \sum \sum \sum_{(j,k,l) \neq} \left(\prod_{\nu \neq j,k,l} \mu_\nu(t) \right) \mathbf{E}\{\tilde{V}_{jkl} \exp(it\hat{\sigma}_N^{-1}(\hat{S}_N^{(j)} + \hat{S}_N^{(k)} + \hat{S}_N^{(l)}))\}, \end{aligned}$$

where $\mu_\nu(t)$ is the characteristic function of the r.v. $\hat{\sigma}_N^{-1}\hat{S}_N^{(\nu)}$. It follows by Lemma XV.4.1 of Feller [7] and (A.5) that

$$(A.7) \quad \mathbf{E}\{\tilde{V}_{jkl} \exp(it\hat{\sigma}_N^{-1}(\hat{S}_N^{(j)} + \hat{S}_N^{(k)} + \hat{S}_N^{(l)}))\} = O(t^2 |c_j^*| \{(c_j^*)^2 + (c_k^*)^2 + (c_l^*)^2 + N^{-1}\omega_2\}).$$

Hence for $r=1$, (A.1) follows from (A.2), (A.6), and (A.7).

We next consider the expansion of $\{\hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\}^r$. Expanding it directly, we find that

$$(A.8) \quad \{\hat{\sigma}_N^{-1}(S_N^0 - \hat{S}_N^0 + I_{22}^0)\}^r = (N+1)^{-2r} \left(\sum_{(j,k,l) \neq} \sum \sum \tilde{V}_{jkl} \right)^r = (N+1)^{-2r} \sum_{b=3}^{3r} \sum_{\nu} q_{b\nu}^{(r)},$$

where $q_b^{(r)}$ is of the form

$$(A.9) \quad q_b^{(r)} = \sum_{(j_1, j_2, \dots, j_b) \neq} \sum \cdots \sum \tilde{V}_{aa_1} \tilde{V}_{aa_2} \cdots \tilde{V}_{aa_r}$$

and each $\langle a \rangle$ is one of the indices j_1, j_2, \dots, j_b . Thus we have

$$(A.10) \quad \begin{aligned} & |\mathbf{E} \exp(it\hat{\sigma}_N^{-1}\hat{S}_N^0) \{ \hat{\sigma}_N^{-1}(\hat{S}_N^0 - \hat{S}_{22}^0 + I_{22}^0) \}^r| \\ &= \left| (N+1)^{-2r} \sum_{b=3}^{3r} \left(\prod_{j \neq j_1, \dots, j_b} \mu_j(t) \right) \sum_{\nu} \mathbf{E} \left\{ q_{b\nu}^{(r)} \prod_{k=1}^b \exp(it\hat{\sigma}_N^{-1}\hat{S}_N^{(j_k)}) \right\} \right| \\ &\leq \exp \left(-\frac{7t^2}{18} + \frac{t^2}{6} r \max_{1 \leq j \leq N} |c_j^*| \right) \sum_{b=3}^{3r} \sum_{\nu} N^{-2r} \left| \mathbf{E} \left\{ q_{b\nu}^{(r)} \prod_{k=1}^b \exp(it\hat{\sigma}_N^{-1}\hat{S}_N^{(j_k)}) \right\} \right| \end{aligned}$$

in view of (A.2).

Let $Q_b^{(r)}$, $3 \leq b \leq 3r$, be the cardinality of the collection of all different terms $q_b^{(r)}$ and put $\Omega_r = \sum_{b=3}^{3r} Q_b^{(r)}$, which is the total number of terms $q_b^{(r)}$ in the expansion (A.8). Then clearly $\Omega_1 = 1$. To estimate Ω_r , we first note the recursive relation

$$\Omega_r \leq \Omega_{r-1} \{ (\alpha^2 - \alpha + 1) + \alpha^2 + \alpha + 1 \} \leq 2\alpha^3 \Omega_{r-1},$$

where $\alpha = 3(r-1)$. Thus there is a constant K_1 (independent of N and r) such that, for any positive integer r ,

$$(A.11) \quad \Omega_r \leq (54)^{r-1} ((r-1)!)^3 \leq K_1^r r^{3r},$$

where the second inequality follows by Stirling's formula (see, e.g., Feller [6]).

To complete the proof, it remains to show that for any positive integer r and b , $3 \leq b \leq 3r$,

$$(A.12) \quad N^{-2r} \left| \mathbf{E} q_b^{(r)} \prod_{k=1}^b \exp(it\hat{\sigma}_N^{-1}\hat{S}_N^{(j_k)}) \right| \leq K_2^r \omega_3 (1 + |t| + t^2)^r,$$

where K_2 is an absolute constant. We shall prove this by induction on r . Note that (A.12) holds for $r=1$ in view of (A.7). We now suppose that (A.12) holds for $r \leq m-1$ and let $r = m \geq 2$.

If $b \leq 2m = 2r$, then (A.12) is trivial. Thus we consider only terms $q_b^{(m)}$, $b \geq 2m+1$. Pick any s , $1 \leq s \leq m$, and let $b = 2m + s$. Let γ denote the number of indices j (in (A.9)) which appear at least twice as a subscript of one of the \tilde{V} -terms and then we have that $3m \geq 2\gamma + 2m + s - \gamma$, i.e.,

$$(A.13) \quad 0 \leq \gamma \leq m - s.$$

Hence at least $m + 2s$ indices among $j_1, j_2, \dots, j_{2m+s}$ appear exactly once. If one of the \tilde{V} -terms has three subscripts appearing only once, then the expectation in (A.12) is split into expectations of two groups to ensure (A.12) immediately by the induction hypothesis. This is the case when $s > m/2$ or $\gamma < s$.

Next we consider the remaining case when

$$(A.14) \quad 1 \leq s \leq m/2, \quad s \leq \gamma \leq m - s$$

and none of the \tilde{V} -terms in (A.9) satisfy the condition that every one of its subscripts appears only once. Since at least $m + 2s$ indices among $j_1, j_2, \dots, j_{2m+s}$ appear exactly once, we must have at least $2s$ \tilde{V} -terms, of which exactly two subscripts appear only once. Thus, by rearranging the order of summations, we can write the typical form of

$q_b^{(r)} \prod_{k=1}^b \exp(it\hat{\sigma}_N^{-1} \hat{S}_N^{(j_k)})$ in the following way:

$$\begin{aligned}
 (A.15) \quad & N^{-2m} \left(\sum_{k_1} \sum_{l_1} \cdots \sum_{k_\alpha} \sum_{l_\alpha} \right) \left(\sum_{m_1} \cdots \sum_{m_\beta} \right) \left(\sum_{j_1} \cdots \sum_{j_\gamma} \right) \\
 & \times \tilde{V}_{ak_1 l_1} \cdots \tilde{V}_{ak_\delta l_\delta} \tilde{V}_{k_{\delta+1} a l_{\delta+1}} \cdots \tilde{V}_{k_\alpha l_\alpha a} \text{ --- (A)} \\
 & \times \tilde{V}_{aam_1} \cdots \tilde{V}_{aam_r} \tilde{V}_{am_{r+1} a} \cdots \tilde{V}_{m\beta aa} \text{ --- (B)} \\
 & \times \tilde{V}_{aaa_1} \cdots \tilde{V}_{aaa_\eta} \text{ --- (C)} \\
 & \times \prod_{\nu=1}^{\alpha} \exp(it\hat{\sigma}_N^{-1}(\hat{S}_N^{(k_\nu)} + \hat{S}_N^{(l_\nu)})) \prod_{\nu=1}^{\beta} \exp(it\hat{\sigma}_N^{-1} \hat{S}_N^{(m_\nu)}) \prod_{\nu=1}^{\gamma} \exp(it\hat{\sigma}_N^{-1} \hat{S}_N^{(j_\nu)}),
 \end{aligned}$$

where summations are taken over distinct indices; indices k , l , and m appear only once; j 's appear at least twice among places marked by $\langle\langle a \rangle\rangle$; $\alpha, \beta, \gamma, \delta, \varepsilon$, and η are non-negative integers such that

$$\begin{aligned}
 (A.16) \quad & 1 \leq s \leq a/2 \leq m/2 \quad s \leq \gamma \leq m-s, \\
 & 0 \leq \delta \leq \alpha \leq m, \quad 0 \leq \varepsilon \leq \beta \leq m, \\
 & 2\alpha + \beta + \gamma = 2m + s, \quad \alpha + \beta + \eta = m.
 \end{aligned}$$

We next consider the conditional expectation of (A.15) given r.v.'s $Z_{j_1}, Z_{j_2}, \dots, Z_{j_\gamma}$. To this end, applying Lemma XV.4.1 of Feller [7] and then taking conditional expectations, we obtain that for distinct indices j , k , and l

$$\begin{aligned}
 (A.17) \quad & \sum_{(k,l) \neq} \sum \mathbf{E}(\tilde{V}_{jkl} \tilde{e}_{kl} | Z_j) \leq \sum_{(k,l) \neq} \sum |t| \{|c_k^*| + |c_l^*| + N^{-1} \omega_1\} \leq 3|t| N^{5/3} \omega_3^{1/3}, \\
 & \sum_{(k,l) \neq} \sum \mathbf{E}(\tilde{V}_{kjl} \tilde{e}_{kl} | Z_j) \leq \sum_{(k,l) \neq} \sum |t| \{|c_k^*| \{|c_k^*| + |c_l^*| + N^{-1} \omega_1\} \} \leq 3|t| N^{4/3} \omega_\beta^{2/3}, \\
 & \sum_{(k,l) \neq} \sum \mathbf{E}(\tilde{V}_{klj} \tilde{e}_{kl} | Z_j) \leq 3|t| N^{4/3} \omega_3^{2/3},
 \end{aligned}$$

where $\tilde{e}_{kl} = \exp(it\hat{\sigma}_N^{-1}(\hat{S}_N^{(k)} + \hat{S}_N^{(l)}))$.

We finally split the proof into two cases when $\alpha = m$ or $\alpha < m$. We first prove (A.12) when $\alpha = m$ and then when $\alpha < m$.

When $\alpha = m$, we have only group (A) in (A.15). If $s \geq 2$, then the expectation of (A.15) is split into two groups and the induction hypothesis ensures (A.12) immediately. If $s = 1$ and $\delta = 0$, then it follows by substitution of (A.17) into (A.15) that the expectation of (A.15) is bounded by

$$N^{-2m} (3|t| N^{4/3} \omega_3^{2/3})^m N \leq (3|t|)^m \omega_3.$$

If $s = 1$ and $\delta > 0$ then the expectation is bounded by

$$N^{-2m} (3|t| N^{5/3} \omega_3^{1/3})^m \omega_1 \leq (3|t|)^m \omega_3.$$

This completes the proof for the case when $\alpha = m$.

We now suppose that $\alpha < m$. If some of the indices j appear only in group (A) of (A.15), then we can again split the expectation into two groups. Thus it remains to show (A.12) when all indices j in group (A) appear again in group (B) or (C). Note that $\beta + \gamma$ is the total number of indices m 's and j 's which appear in group (B) or (C) and that the number of indices m 's and j 's, which appear as one of the second or the third subscripts of the \tilde{V} -terms in group (B) or (C), is at most $2(\eta + \beta)$. Moreover,

$$\beta + \gamma - 2(\eta + \beta) = \gamma - 2(m - \alpha - \beta) - \beta = s$$

in view of (A.16). Thus at least s indices among $m_1, m_2, \dots, m_\beta, j_1, j_2, \dots, j_\gamma$ must appear as the first subscript of one of the \tilde{V} -terms in group (B) or (C). Utilizing this

fact and substituting (A.17) into (A.15), we find that the expectation of (A.15) is bounded by

$$\begin{aligned} & N^{-2m} (3|t|N^{5/3}\omega_3^{1/3})^\delta (3|t|N^{4/3}\omega_3^{2/3})^{\alpha-\delta} \omega_1^s N^{2m+s-(2\alpha+s)} \\ & \leq (3|t|)^\alpha N^{-2\alpha+\delta/3+4\alpha/3+2s/3} \omega_3^{2\alpha/3-\delta/3+s/3} \\ & \leq (3(1+|t|))^m N^{-(\alpha-2s)/3} \omega_3^{(\alpha+s)/3} \leq (3(1+|t|))^m \omega_3 \end{aligned}$$

because $2s \leq \alpha$ and $\delta \leq \alpha$. This completes the proof of the case when $\alpha < m$. The proof follows.

REFERENCES

- [1] H. BERGSTRÖM AND M. L. PURI, *Convergence and remainder terms in linear rank statistics*, Ann. Statist., 5 (1977), pp. 671-680.
- [2] P. K. BHATTACHARYA, *Convergence of sample path of normalized sums of induced order statistics*, Ann. Statist., 2 (1974), pp. 1034-1039.
- [3] S. BJERVE, *Error bounds for linear combinations of order statistics*, Ann. Statist., 5 (1977), pp. 357-369.
- [4] R. J. M. M. DOES, *Berry-Esseen theorems for simple linear rank statistics under null-hypothesis*, Ann. Probab., 10 (1982), pp. 982-991.
- [5] W. FELLER, *Generalization of a probability limit theorem of Cramér*, Trans. Amer. Math. Soc., 54 (1943), pp. 361-372.
- [6] ———, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed., John Wiley, New York, 1968.
- [7] ———, *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd ed., John Wiley, New York, 1971.
- [8] J. HÁJEK, *Asymptotically most powerful rank order tests*, Ann. Math. Statist., 33 (1962), pp. 1129-1147.
- [9] ———, *Asymptotic normality of simple linear rank statistics under alternatives*, Ann. Math. Statist., 39 (1968), pp. 325-346.
- [10] R. HELMERS, *The order of the approximation for linear combinations of order statistics with smooth weight functions*, Ann. Probab., 5 (1977), pp. 940-953.
- [11] ———, *A Berry-Esseen theorem for linear combinations of order statistics*, Ann. Probab., 9 (1981), pp. 342-347.
- [12] W. HOEFFDING, *A combinatorial central limit theorem*, Ann. Math. Statist., 22 (1951), pp. 558-566.
- [13] M. HUŠKOVÁ, *Asymptotic distribution of simple linear rank statistics for testing symmetry*, Z. Wahrsch. Verw. Gebiete, 14 (1970), pp. 308-322.
- [14] ———, *The rate of convergence of simple linear rank statistics under hypothesis and alternatives*, Ann. Statist., 5 (1977), pp. 658-670.
- [15] ———, *The Berry-Esseen theorem for rank statistics*, Comment. Math. Univ. Carolin., 20 (1979), pp. 399-415.
- [16] J. JUREČKOVÁ AND M. L. PURI, *Order of normal approximation for rank test statistics distribution*, Ann. Probab., 3 (1975), pp. 526-533.
- [17] W. C. M. KALLENBERG, *Cramér type large deviations for simple linear rank statistics*, Z. Wahrsch. Verw. Gebiete, 60 (1982), pp. 403-409.
- [18] M. MOTOO, *On the Hoeffding's combinatorial central limit theorem*, Ann. Inst. Statist. Math., 8 (1957), pp. 145-154.
- [19] U. MÜLLER-FUNK AND H. WITTING, *On the rate of convergence in CLT for signed linear rank statistics*, Nonparametric Statistical Inference, Vol. II, B. V. Gnedenko, M. L. Puri, and I. Vincze, eds., Elsevier, Amsterdam, 1982, pp. 637-652.
- [20] V. V. PETROV, *Sums of Independent Random Variables*, Springer-Verlag, Berlin, New York, 1975.
- [21] S. S. RALESCU AND M. L. PURI, *On the rate of convergence in the central theorem for signed rank statistics*, Adv. Appl. Math., (1985), pp. 23-51.
- [22] S. M. STIGLER, *Linear functions of order statistics with smooth weight functions*, Ann. Statist., 2 (1974), pp. 676-693.
- [23] P. VAN BEEK, *An application of Fourier methods to the problem of sharpening the Berry-Esseen inequality*, Z. Wahrsch. Verw. Gebiete, 23 (1972), pp. 187-196.
- [24] M. VANDERMAELE AND N. VERAVERBEKE, *Cramér-type large deviations and moderate deviations for linear combinations of order statistics*, Ann. Probab., 10 (1982), pp. 423-434.

- [25] V. VON BAHR, *Remainder term estimate in a combinatorial limit theorem*, Z. Wahrsch. Verw. Gebiete, 35 (1976), pp. 131–139.
- [26] S. YANG, *Linear functions of concomitants of order statistics with application to nonparametric estimate of a regression function*, J. Amer. Statist. Assoc., 76 (1981), pp. 658–662.
- [27] M. L. PURI AND M. SEOH, *On the rate of convergence to normality for generalized linear rank statistics*, Ann. Inst. Statist. Math., 37 (1985), pp. 51–67.
- [28] M. L. PURI AND S. S. RALESCU, *Almost sure linearity for signed rank statistics in the non-i.i.d. case*, Acta Math. Hungar., 48 (1986), pp. 273–284.
- [29] M. SEOH, S. S. RALESCU, AND M. L. PURI, *Cramér-type large deviations for generalized rank statistics*, Ann. Probab., 13 (1985), pp. 115–125.
- [30] M. SEOH AND M. L. PURI, *Asymptotic normality of a class of nonparametric statistics*, Econ. Theory, 3 (1987), pp. 313–347.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.